# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| 3/31/2005 | Final Report | 9/15/02 - 12/31/04 |

**4. TITLE AND SUBTITLE**
Genomics of the Human Gastrointestinal Microbiome

**5a. CONTRACT NUMBER**
N00014-02-1-1002

**5b. GRANT NUMBER**
N00014-02-1-1002

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Steven R. Gill

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Institute for Genomic Research
9712 Medical Center Drive
Rockville, MD 20850

**8. PERFORMING ORGANIZATION REPORT NUMBER**
N00014-02-1-1002

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Office of Naval Research Regional Office Atlanta
100 Alabama Street, SW Suite 4R15
Atlanta, GA 30303-3104

**10. SPONSOR/MONITOR'S ACRONYM(S)**
ONR

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**
N00014-02-1-1002

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
"Approved for Public Release, Distribution Unlimited"

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

We have used sequencing of targeted 16S rDNA PCR libraries and random metagenomic libraries to examine the phylogenetic and genomic diversity within the human gastrointestinal bacterial community. In one study, we used both 16S and metagenomic analysis in a study of fecal samples from three healthy human subjects. In a second study, we used 16S analysis to compare bacterial diversity between samples obtained from multiple intestinal mucosal sites and companion fecal samples of multiple healthy individuals. These studies demonstrate a remarkable diversity between bacterial communities in the different sites within the gastrointestinal tract and also between individuals. Through the metagenomic approach, we have identified genome sequences from fungal and archael organisms and identified several virulence genes as potential markers of pathogens within the population. Finally, we have modified our DNA assembly program, Celera Assembler, to successfully assemble contigs up to ~40 kb from the random metagenomic sequence data.

**15. SUBJECT TERMS**
Human gastrointestinal bacterial community, 16S rDNA, metagenomics

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Leslie Bisignano |
| Unclass. | Unclass. | Unclass. | UL | | 19b. TELEPHONE NUMBER *(Include area code)* (301) 795-7591 |

**GRANT #:** ONR-N00014-02-1-1002

**PRINCIPAL INVESTIGATOR:** Dr. Steven R. Gill

**INSTITUTION:** The Institute for Genomic Research

**GRANT TITLE:** Genomics of the Human Gastrointestinal Microbiome

**AWARD PERIOD:** 09/15/02-12/31/04

**ABSTRACT:**

We have used sequencing of targeted 16S rDNA PCR libraries and random metagenomic libraries to examine the phylogenetic and genomic diversity within the human gastrointestinal bacterial community. In one study, we used both 16S and metagenomic analysis in a study of fecal samples from three healthy human subjects. In a second study, we used 16S analysis to compare bacterial diversity between samples obtained from multiple intestinal mucosal sites and companion fecal samples of multiple healthy individuals. These studies demonstrate a remarkable diversity between bacterial communities in the different sites within the gastrointestinal tract and also between individuals. Through the metagenomic approach, we have identified genome sequences from fungal and archaeal organisms and identified several virulence genes as potential markers of pathogens within the population. Finally, we have modified our DNA assembly program, Celera Assembler, to successfully assemble contigs up to ~40 kb from the random metagenomic sequence data.

**KEY WORDS:**

Human gastrointestinal bacterial community, 16S rDNA, metagenomics

**OBJECTIVE:**

Humans are colonized with a dynamic and intricate bacterial community that covers the surfaces of the skin, oral cavity and gastrointestinal tract. This bacterial community and the human host co-exist in a mutually beneficial or commensal relationship, with the bacteria playing a pivotal role in human development and evolution. Bacterial density is highest in the gastrointestinal tract, where the bacterial population outnumbers surrounding host cells by an order of magnitude and accounts for as much as 50% of fecal material dry weight. Intimate interactions of the bacterial communities with intestinal mucosal surfaces contribute to both innate and adaptive immunity and protection against epithelial cell injury. Disruption of these interactions frequently leads to immune disorders such as inflammatory bowel diseases. In terms of impact on human nutrition, co-evolution of this bacterial community with the human host has led to development of bacterial metabolic activities capable of extracting maximum nutritional value from food ingested by the host and on regulation of host fat storage.

In spite of current progress, our understanding of the gastrointestinal microbial communities, and of all human borne microbial communities in general, is limited by our inability to cultivate a majority of these bacteria, and our limited understanding of shifts in human microbial populations occurring as a result of differences in age, sex or diet of the human host.

The work outlined in our proposal was designed to explore the feasibility of using targeted 16S rDNA and random metagenomic sequencing to identify members of the gastrointestinal microbiome and to characterize the genomic diversity of this population. This inventory will facilitate clinical and basic research focused towards understanding the metabolic potential of the gastrointestinal microbiome and the exquisite balance between a healthy, symbiotic relationship and one of disease. An understanding of these disease dynamics will also impact formation of strategies to counter threats of biologic warfare.

**APPROACH:**

We isolated bacterial genomic DNA from multiple human intestinal mucosal surfaces and companion fecal samples. The DNA was then either PCR amplified with broad-range bacterial or archaeal 16S rDNA primers or used to construct random metagenomic sequencing libraries (average insert size ~2 kb). The PCR products were sequenced, assembled into ~1.5 kb contigs, and analyzed by comparison to the ribosomal DNA sequence database in Ribosomal Database Project II. The random metagenomic libraries were sequenced (~150,000 total sequences) and assembled into contigs using a modified Celera Assembler program. Genes within these contigs, as well as genes outside of the contigs, were analyzed by blastx to identify functional proteins and determine the number of novel genes present in our samples.

**ACCOMPLISHMENTS:**

We have successfully used a metagenomics approach to examine genomic diversity of the human intestinal bacterial community. In this effort, we identified novel members of the intestinal community and ~20,000 previously unidentified genes. We have also optimized the Celera Assembly program to assemble the random sequence data from the metagenomic libraries. In a collaborative effort with David Relman, we have 16S rDNA analysis to demonstrate remarkable diversity of the bacterial community at different locations within the intestinal tract. A manuscript describing the results from the 16S work has been accepted for publication in Science. A manuscript describing the metagenomic effort is currently being prepared for submission.

## CONCLUSIONS:

*Conclusions from the 16S rDNA effort on intestinal mucosal and fecal samples.* This work has shown that bacterial diversity within the human intestinal tract is greater than previously described and that much of this diversity is novel. Differences in diversity between individuals was significantly greater than intra-subject differences. Differences in fecal diversity are much greater than the variation among subject-specific fecal libraries. One conclusion from this work is that the fecal bacterial community represents a combination of shed mucosal bacteria and a separate non-adherent luminal population.

*Conclusions from the metagenomic effort on fecal samples.* This work supports results from the 16S rDNA analysis showing that there is an unexpected level of genetic/genomic diversity in the fecal bacterial community. In addition to diversity within the bacterial community, a significant number of archaeal and fungal organisms were identified in our study. The most surprising result of our study was the identification of possible genetic strain variation among multiple *Bifidobacterium longum* genes, indicating highly variable groups of this key intestinal bacterium. Finally, the high number of novel genes present in our samples indicates possible novel bacteria in the intestinal community.

## SIGNIFICANCE:

The human intestinal bacterial community plays an essential role in human development and evolution. A complete definition of the members of this community and their genomic diversity will lay the groundwork for development of treatments for diseases such as inflammatory bowel disease and methods for maintaining a healthy intestinal ecosystem. This information is also essential for the development of strategies to counter threats of biological warfare which use bacterial agents for induction of severe emetic responses and dehydration. Our data illustrates the high level of diversity present in these bacterial populations and has given us a reference point for the genera and species present in a healthy human. Further studies are needed to explore the impact of host genetics and of perturbations such as antimicrobials and change in diet.

## PATENT INFORMATION: none

## AWARD INFORMATION: none

## PUBLICATIONS AND ABSTRACTS:

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Sargent, M., Dethlefsen, L., Purdom, E., **Gill, S.R.**, Nelson, K.E. and Relman, D.A. (2005) Diversity of the human intestinal microflora at multiple anatomical sites using large-scale 16S rDNA sequence analysis. In press, April 2005, *Science*.

**Gill, S.R.**, Deboy, R., Hance, I., Relman, D.A., Nelson, K.E. (2005) Diversity within the human gastrointestinal microbiome. *In prep*.

**Gill, S.R.**, Deboy, B., Hance, I., Relman, D.A., Nelson, K.E. (2005) Genomics of the Human Gastrointestinal Microbiome. American Society for Microbiology Conference on Beneficial Microbes.